Computers and Mathematics with Applications 63 (2012) 975–984

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Two-stage least squares based iterative identification algorithm for controlled autoregressive moving average (CARMA) systems[☆]

Guoyu Yao[a], Ruifeng Ding[b,*]

[a] Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, PR China

[b] School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, PR China

ARTICLE INFO

Article history:
Received 12 July 2011
Received in revised form 1 December 2011
Accepted 1 December 2011

Keywords:
Least squares
Parameter estimation
Iterative algorithm
Two-stage identification

ABSTRACT

A two-stage least squares based iterative (two-stage LSI) identification algorithm is derived for controlled autoregressive moving average (CARMA) systems. The basic idea is to decompose a CARMA system into two subsystems and to identify each subsystem, respectively. Because the dimensions of the involved covariance matrices in each subsystem become small, the proposed algorithm has a high computational efficiency. The simulation results indicate that the proposed algorithm is effective.

1. Introduction

The iterative numerical algorithms can be used to solve not only some matrix equations [1–9], but also parameter estimation and filtering problems [10–18]. The iterative identification algorithms make sufficient use of all input–output data and can improve the parameter estimation accuracy [19–21]. For example, Ding et al. presented least squares based iterative algorithms for Hammerstein nonlinear ARMAX systems [22], and for OE and OEMA systems [19]; Han et al. gave a hierarchical least squares based iterative identification algorithm for a class of multivariable CARMA-like systems [23]. Zhang et al. proposed a hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems (i.e., multivariable OEMA-like systems) [24]. Bao et al. developed a least squares based iterative identification method for multivariable controlled ARMA systems [25]. Ding et al. presented a least squares based iterative algorithm for controlled autoregressive autoregressive moving average (CARARMA) systems [26]. Also, a least squares based iterative algorithm and a gradient based iterative algorithm are developed for Box–Jenkins systems [20,21].

Two-stage algorithms have been widely used in the identification field. Bai presented an optimal two-stage identification algorithm for Hammerstein–Wiener nonlinear systems in the sense of a weighted nonlinear least squares criterion [27]. Li et al. gave a two-stage algorithm identification of nonlinear dynamic systems to reduce computational complexity [28]. Hwang et al. developed a two-stage least squares algorithm to identify continuous systems with time delay based on the pulse responses [29]. Cao gave a rounding error analysis of two-stage iterative methods for large linear systems [30].

The least squares based iterative algorithm was used to identify the Box–Jenkins models with finite measurement data [21], but it requires large computational load, especially when the dimension of the involved covariance matrices is large. Duan et al. presented a two-stage recursive least squares parameter estimation algorithm for output error models [31]. This paper uses the decomposition technique and proposes a two-stage iterative identification method for controlled

☆ This work was supported in part by the National Natural Science Foundation of China.

* Corresponding author.
E-mail addresses: yaoguoyu123@126.com (G. Yao), rfding@yahoo.cn (R. Ding).

0898-1221/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2011.12.002

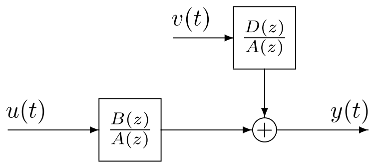

Fig. 1. A system described by the CARMA model.

autoregressive moving average (CARMA) systems. The basic idea is to decompose a CARMA system into two subsystems, to transform the original identification problem into two subproblems with small dimensions, and to identify the parameters of each subsystem, respectively. The proposed algorithm has a high computational efficiency and can be extended to nonlinear systems [32,33].

For several decades, many recursive and iterative parameter estimation algorithms have been developed, including the least squares-based iterative estimation for output error moving average (OEMA) systems using data filtering [34], the input–output data filtering based recursive least squares parameter estimation for CARARMA systems [35], the auxiliary model based least squares or gradient estimation algorithms [36–42], the multi-innovation parameter estimation algorithm for linear and pseudo-linear regression systems [43–52], the hierarchical least squares or hierarchical gradient parameter estimation algorithms [53–63], and other parameter estimation algorithms for linear regressive models [64–73], the recursive and iterative identification methods for Hammerstein systems [74], the identification algorithms for Hammerstein OEMA systems, Hammerstein OEAR systems and Wiener systems [75–80].

The rest of the paper is organized as follows. Section 2 derives a two-stage least squares based iterative identification algorithm for CARMA systems. Section 3 gives the least squares based iterative algorithm for comparisons. Section 4 provides a simulation example to illustrate the effectiveness of the proposed algorithms. Finally, we offer some concluding remarks in Section 5.

2. The two-stage least squares algorithm

Let us introduce some notations first. "$\boldsymbol{A} =: \boldsymbol{X}$" or "$\boldsymbol{X} := \boldsymbol{A}$" stands for "$\boldsymbol{A}$ is defined as $\boldsymbol{X}$"; the symbol $\boldsymbol{I}$ ($\boldsymbol{I}_n$) stands for an identity matrix of appropriate size ($n \times n$); $\mathbf{1}_n$ represents an n-dimensional column vector whose elements are all 1; the superscript T denotes the matrix transpose; the norm of a matrix $\boldsymbol{X}$ is defined by $\|\boldsymbol{X}\|^2 = \mathrm{tr}[\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}]$.

Consider the CARMA system, depicted in Fig. 1,

$$A(z)y(t) = B(z)u(t) + D(z)v(t), \tag{1}$$

where $\{u(t)\}$ and $\{y(t)\}$ are the input and output sequences of the system, respectively, $\{v(t)\}$ is a white noise sequence with zero mean and variance σ^2, and $A(z)$, $B(z)$ and $D(z)$ are the polynomials, of known orders (n_a, n_b, n_d), in the unit backward shift operator z^{-1} [i.e., $z^{-1}y(t) = y(t-1)$], and defined by

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a},$$
$$B(z) := b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b},$$
$$D(z) := 1 + d_1 z^{-1} + d_2 z^{-2} + \cdots + d_{n_d} z^{-n_d}.$$

Without loss of generality, assume that $y(t) = 0$, $u(t) = 0$ and $v(t) = 0$ for $t \leqslant 0$.

The objective of this paper is to apply the decomposition technique and derive a two-stage least squares based iterative identification algorithm for estimating the system parameters a_i, b_i and d_i.

Define the system parameter vectors,

$$\boldsymbol{\Theta} := \begin{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{\vartheta} \end{bmatrix} \in \mathbb{R}^n, \quad n := n_a + n_b + n_d,$$
$$\boldsymbol{\theta} := [a_1, a_2, \ldots, a_{n_a}, b_1, b_2, \ldots, b_{n_b}]^{\mathrm{T}} \in \mathbb{R}^{n_a+n_b},$$
$$\boldsymbol{\vartheta} := [d_1, d_2, \ldots, d_{n_d}]^{\mathrm{T}} \in \mathbb{R}^{n_d},$$

and the information vectors,

$$\boldsymbol{\varphi}(t) := \begin{bmatrix} \boldsymbol{\phi}(t) \\ \boldsymbol{\psi}(t) \end{bmatrix} \in \mathbb{R}^n,$$

$$\boldsymbol{\phi}(t) := [-y(t-1), -y(t-2), \ldots, -y(t-n_a), u(t-1), u(t-2), \ldots, u(t-n_b)]^{\mathrm{T}} \in \mathbb{R}^{n_a+n_b}, \tag{2}$$

$$\boldsymbol{\psi}(t) := [v(t-1), v(t-2), \ldots, v(t-n_d)]^{\mathrm{T}} \in \mathbb{R}^{n_d}. \tag{3}$$

From (1), we can obtain the following identification model:

$$y(t) = \boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\theta} + \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta} + v(t) \tag{4}$$
$$= \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\Theta} + v(t). \tag{5}$$

Define two intermediate variables:

$$y_1(t) := y(t) - \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}, \tag{6}$$

$$y_2(t) := y(t) - \boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\theta}. \tag{7}$$

From (6) and (7), the system in (4) can be decomposed into two "fictitious" subsystems:

$$y_1(t) = \boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t), \tag{8}$$

$$y_2(t) = \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta} + v(t). \tag{9}$$

Consider the data from $t = 1$ to $t = L$ $(L \gg n)$ and define the stacked output vectors $\boldsymbol{Y}(L)$, $\boldsymbol{Y}_1(L)$ and $\boldsymbol{Y}_2(L)$, the stacked information matrices $\boldsymbol{\Phi}(L)$ and $\boldsymbol{\Psi}(L)$ and the stacked white noise vector $\boldsymbol{V}(L)$ as

$$\boldsymbol{Y}(L) := \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(L) \end{bmatrix} \in \mathbb{R}^L, \qquad \boldsymbol{V}(L) := \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(L) \end{bmatrix} \in \mathbb{R}^L,$$

$$\boldsymbol{Y}_1(L) := \begin{bmatrix} y_1(1) \\ y_1(2) \\ \vdots \\ y_1(L) \end{bmatrix} \in \mathbb{R}^L, \qquad \boldsymbol{Y}_2(L) := \begin{bmatrix} y_2(1) \\ y_2(2) \\ \vdots \\ y_2(L) \end{bmatrix} \in \mathbb{R}^L,$$

$$\boldsymbol{\Phi}(L) := \begin{bmatrix} \boldsymbol{\phi}^{\mathrm{T}}(1) \\ \boldsymbol{\phi}^{\mathrm{T}}(2) \\ \vdots \\ \boldsymbol{\phi}^{\mathrm{T}}(L) \end{bmatrix} \in \mathbb{R}^{L \times (n_a + n_b)}, \qquad \boldsymbol{\Psi}(L) := \begin{bmatrix} \boldsymbol{\psi}^{\mathrm{T}}(1) \\ \boldsymbol{\psi}^{\mathrm{T}}(2) \\ \vdots \\ \boldsymbol{\psi}^{\mathrm{T}}(L) \end{bmatrix} \in \mathbb{R}^{L \times n_d}.$$

Note that $\boldsymbol{Y}(L)$, $\boldsymbol{\Phi}(L)$ and $\boldsymbol{\Psi}(L)$ contain all the measured data $\{u(t), y(t) : t = 1, 2, \ldots, L\}$. From (6) and (7), we have

$$\boldsymbol{Y}_1(L) = \boldsymbol{Y}(L) - \boldsymbol{\Psi}(L)\boldsymbol{\vartheta}, \tag{10}$$

$$\boldsymbol{Y}_2(L) = \boldsymbol{Y}(L) - \boldsymbol{\Phi}(L)\boldsymbol{\theta}. \tag{11}$$

From (8) and (9), we have

$$\boldsymbol{Y}_1(L) = \boldsymbol{\Phi}(L)\boldsymbol{\theta} + \boldsymbol{V}(L),$$

$$\boldsymbol{Y}_2(L) = \boldsymbol{\Psi}(L)\boldsymbol{\vartheta} + \boldsymbol{V}(L).$$

Define two quadratic criterion functions:

$$J_1(\boldsymbol{\theta}) := \|\boldsymbol{Y}_1(L) - \boldsymbol{\Phi}(L)\boldsymbol{\theta}\|^2,$$

$$J_2(\boldsymbol{\vartheta}) := \|\boldsymbol{Y}_2(L) - \boldsymbol{\Psi}(L)\boldsymbol{\vartheta}\|^2.$$

For these two optimization problems, letting the partial derivatives of $J_1(\boldsymbol{\theta})$ and $J_2(\boldsymbol{\vartheta})$ with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$ be zero gives

$$\frac{\partial J_1(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = -2\boldsymbol{\Phi}^{\mathrm{T}}(L)[\boldsymbol{Y}_1(L) - \boldsymbol{\Phi}(L)\boldsymbol{\theta}] = \mathbf{0},$$

$$\frac{\partial J_2(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = -2\boldsymbol{\Psi}^{\mathrm{T}}(L)[\boldsymbol{Y}_2(L) - \boldsymbol{\Psi}(L)\boldsymbol{\vartheta}] = \mathbf{0}.$$

Assume that the information vectors $\boldsymbol{\phi}(t)$ and $\boldsymbol{\psi}(t)$ are persistently exciting, that is, $[\boldsymbol{\Phi}^{\mathrm{T}}(L)\boldsymbol{\Phi}(L)]$ and $[\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{\Psi}(L)]$ are non-singular. From the above two equations, we can obtain the following least squares estimates of the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$:

$$\hat{\boldsymbol{\theta}} = [\boldsymbol{\Phi}^{\mathrm{T}}(L)\boldsymbol{\Phi}(L)]^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(L)\boldsymbol{Y}_1(L), \tag{12}$$

$$\hat{\boldsymbol{\vartheta}} = [\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{\Psi}(L)]^{-1}\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{Y}_2(L). \tag{13}$$

Substituting (10) into (12) and (11) into (13) gives

$$\hat{\boldsymbol{\theta}} = [\boldsymbol{\Phi}^{\mathrm{T}}(L)\boldsymbol{\Phi}(L)]^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(L)[\boldsymbol{Y}(L) - \boldsymbol{\Psi}(L)\boldsymbol{\vartheta}], \tag{14}$$

$$\hat{\boldsymbol{\vartheta}} = [\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{\Psi}(L)]^{-1}\boldsymbol{\Psi}^{\mathrm{T}}(L)[\boldsymbol{Y}(L) - \boldsymbol{\Phi}(L)\boldsymbol{\theta}]. \tag{15}$$

However, the right-hand sides of (14) and (15) contain the unknown parameter $\boldsymbol{\vartheta}$ and $\boldsymbol{\theta}$, respectively, it is impossible to compute the estimates $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\vartheta}}$. The approach is based on the hierarchical identification principle: let $k = 1, 2, 3, \ldots$ be an iteration variable, $\hat{\mathbf{\Theta}}_k := \begin{bmatrix}\hat{\boldsymbol{\theta}}_k\\ \hat{\boldsymbol{\vartheta}}_k\end{bmatrix}$ be the iterative estimate of $\mathbf{\Theta} = \begin{bmatrix}\boldsymbol{\theta}\\ \boldsymbol{\vartheta}\end{bmatrix}$ at iteration k, and $\hat{v}_k(t)$ be the estimate of $v(t)$ at iteration k, and define

$$\hat{\boldsymbol{\varphi}}_k(t) := \begin{bmatrix}\boldsymbol{\phi}(t)\\ \hat{\boldsymbol{\psi}}_k(t)\end{bmatrix} \in \mathbb{R}^{n_a+n_b+n_d},$$

$$\hat{\boldsymbol{\psi}}_k(t) := [\hat{v}_{k-1}(t-1), \hat{v}_{k-1}(t-2), \ldots, \hat{v}_{k-1}(t-n_d)]^{\mathrm{T}} \in \mathbb{R}^{n_d},$$

$$\hat{\mathbf{\Psi}}_k(L) := \begin{bmatrix}\hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(1)\\ \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(2)\\ \vdots\\ \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(L)\end{bmatrix} \in \mathbb{R}^{L\times n_d}.$$

From (4), we have

$$v(t) = y(t) - \boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\theta} - \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}.$$

Replacing $\boldsymbol{\psi}(t)$, $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$ with $\hat{\boldsymbol{\psi}}_k(t)$, $\hat{\boldsymbol{\theta}}_k$ and $\hat{\boldsymbol{\vartheta}}_k$, the estimate $\hat{v}_k(t)$ of $v(t)$ can be computed by

$$\begin{aligned}\hat{v}_k(t) &= y(t) - \boldsymbol{\phi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}_k.\\ &= y(t) - \hat{\boldsymbol{\varphi}}_k^{\mathrm{T}}(t)\hat{\mathbf{\Theta}}_k. \end{aligned} \tag{16}$$

Replacing $\mathbf{\Psi}(L)$ and $\boldsymbol{\vartheta}$ in (14) with their estimates $\hat{\mathbf{\Psi}}_k(L)$ and $\hat{\boldsymbol{\vartheta}}_{k-1}$, and replacing $\mathbf{\Psi}(L)$ and $\boldsymbol{\theta}$ in (15) with their estimates $\hat{\mathbf{\Psi}}_k(L)$ and $\hat{\boldsymbol{\theta}}_k$, we can summarize the two-stage least squares based iterative (two-stage LSI) identification algorithm for estimating $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$ of the CARMA systems as follows:

$$\hat{\boldsymbol{\theta}}_k = [\mathbf{\Phi}^{\mathrm{T}}(L)\mathbf{\Phi}(L)]^{-1}\mathbf{\Phi}^{\mathrm{T}}(L)[\boldsymbol{Y}(L) - \hat{\mathbf{\Psi}}_k(L)\hat{\boldsymbol{\vartheta}}_{k-1}], \tag{17}$$

$$\hat{\boldsymbol{\vartheta}}_k = [\hat{\mathbf{\Psi}}_k^{\mathrm{T}}(L)\hat{\mathbf{\Psi}}_k(L)]^{-1}\hat{\mathbf{\Psi}}_k^{\mathrm{T}}(L)[\boldsymbol{Y}(L) - \mathbf{\Phi}(L)\hat{\boldsymbol{\theta}}_k], \quad k = 1, 2, \ldots, \tag{18}$$

$$\boldsymbol{Y}(L) = [y(1), y(2), \ldots, y(L)]^{\mathrm{T}}, \tag{19}$$

$$\mathbf{\Phi}(L) = [\boldsymbol{\phi}(1), \boldsymbol{\phi}(2), \ldots, \boldsymbol{\phi}(L)]^{\mathrm{T}}, \tag{20}$$

$$\hat{\mathbf{\Psi}}_k(L) = [\hat{\boldsymbol{\psi}}_k(1), \hat{\boldsymbol{\psi}}_k(2), \ldots, \hat{\boldsymbol{\psi}}_k(L)]^{\mathrm{T}}, \tag{21}$$

$$\boldsymbol{\phi}(t) = [-y(t-1), -y(t-2), \ldots, -y(t-n_a), u(t-1), u(t-2), \ldots, u(t-n_b)]^{\mathrm{T}}, \tag{22}$$

$$\hat{\boldsymbol{\psi}}_k(t) = [\hat{v}_{k-1}(t-1), \hat{v}_{k-1}(t-2), \ldots, \hat{v}_{k-1}(t-n_d)]^{\mathrm{T}}, \tag{23}$$

$$\hat{\boldsymbol{\varphi}}_k(t) = [\boldsymbol{\phi}^{\mathrm{T}}(t), \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{24}$$

$$\hat{\mathbf{\Theta}}_k := [\hat{\boldsymbol{\theta}}_k^{\mathrm{T}}, \hat{\boldsymbol{\vartheta}}_k^{\mathrm{T}}]^{\mathrm{T}}, \tag{25}$$

$$\hat{v}_k(t) = y(t) - \hat{\boldsymbol{\varphi}}_k^{\mathrm{T}}(t)\hat{\mathbf{\Theta}}_k. \tag{26}$$

In the above algorithm, the dimension of the covariance matrix $S_1 := [\mathbf{\Phi}^{\mathrm{T}}(L)\mathbf{\Phi}(L)]^{-1}$ in (17) is $(n_a + n_b) \times (n_a + n_b)$, $S_2 := [\hat{\mathbf{\Psi}}_k^{\mathrm{T}}(L)\hat{\mathbf{\Psi}}_k(L)]^{-1}$ in (18) is $n_d \times n_d$.

The steps involved in the two-stage LSI algorithm in (17)–(26) to compute the parameter estimates $\hat{\boldsymbol{\theta}}_k$ and $\hat{\boldsymbol{\vartheta}}_k$ for CARMA systems are listed below.

1. Collect the input–output data $\{u(t), y(t) : i = 1, 2, \ldots, L\}$ (L is the data length), form $\boldsymbol{Y}(L)$ by (19), $\boldsymbol{\phi}(t)$ by (22) and $\mathbf{\Phi}(L)$ by (20), and give the parameter estimation precision $\varepsilon = 0.01$.
2. To initialize, let $k = 1$, $\hat{\boldsymbol{\vartheta}}_0 = \mathbf{1}_{n_d}/p_0$, $p_0 = 10^6$, $\hat{v}_0(t) =$ a random number.
3. Form $\hat{\boldsymbol{\psi}}_k(t)$ by (23), $\hat{\mathbf{\Psi}}_k(L)$ by (21) and $\hat{\boldsymbol{\varphi}}_k(t)$ by (24).
4. Update $\hat{\boldsymbol{\theta}}_k$ by (17) and $\hat{\boldsymbol{\vartheta}}_k$ by (18).
5. Form $\hat{\mathbf{\Theta}}_k$ by (25).
6. Compute $\hat{v}_k(t)$ by (26).
7. If $\|\hat{\mathbf{\Theta}}_k - \mathbf{\Theta}\| = \|\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-1}\| + \|\hat{\boldsymbol{\vartheta}}_k - \hat{\boldsymbol{\vartheta}}_{k-1}\| \leqslant \varepsilon$, obtain the parameter estimates $\hat{\boldsymbol{\theta}}_k$ and $\hat{\boldsymbol{\vartheta}}_k$; otherwise, increase k by 1 and go to step 3.

The flowchart of computing the parameter estimates $\hat{\boldsymbol{\theta}}_k$ and $\hat{\boldsymbol{\vartheta}}_k$ is shown in Fig. 2.

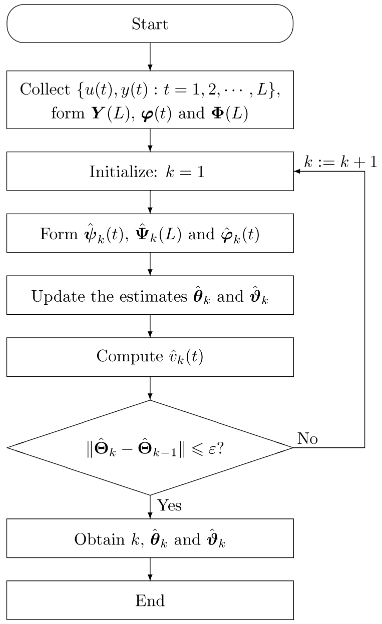

Fig. 2. The flowchart for computing the two-stage LSI parameter estimates.

3. The least squares based iterative algorithm

To show the advantages of the proposed two-stage LSI algorithm, the following gives the least squares based iterative (LSI) algorithm for comparisons. Define the stacked information matrix

$$\boldsymbol{H}(L) := \begin{bmatrix} \boldsymbol{\varphi}^{\mathrm{T}}(1) \\ \boldsymbol{\varphi}^{\mathrm{T}}(2) \\ \vdots \\ \boldsymbol{\varphi}^{\mathrm{T}}(L) \end{bmatrix} \in \mathbb{R}^{L\times(n_a+n_b+n_d)}.$$

Using the definitions of $\boldsymbol{Y}(L)$, $\boldsymbol{V}(L)$ and $\boldsymbol{H}(L)$, Eq. (5) can be written as

$$\boldsymbol{Y}(L) = \boldsymbol{H}(L)\boldsymbol{\Theta} + \boldsymbol{V}(L).$$

Define a quadratic criterion function:

$$J_3(\boldsymbol{\Theta}) := \|\boldsymbol{Y}(L) - \boldsymbol{H}(L)\boldsymbol{\Theta}\|^2.$$

Minimizing the criterion function $J_3(\boldsymbol{\Theta})$ and letting its partial derivative with respect to $\boldsymbol{\Theta}$ be zero, we can obtain the least squares algorithm for estimating the parameter vector $\boldsymbol{\Theta}$:

$$\hat{\boldsymbol{\Theta}} = [\boldsymbol{H}^{\mathrm{T}}(L)\boldsymbol{H}(L)]^{-1}\boldsymbol{H}^{\mathrm{T}}(L)\boldsymbol{Y}(L). \tag{27}$$

Here, a difficulty arises because $\boldsymbol{H}(L)$ in the above equation (that is $\boldsymbol{\varphi}(t)$ in (3)) contains the unknown noise terms $v(t-i)$. The approach is based on the hierarchical identification principle: let $\hat{\boldsymbol{\varphi}}_k(t)$ denote the information vector by replacing $v(t-i)$ in (3) with their estimates $\hat{v}_{k-1}(t-i)$ at iteration $k-1$, $\hat{\boldsymbol{H}}_k(L)$ denote the stacked information matrix obtained by replacing $\boldsymbol{\varphi}(t-i)$ in $\boldsymbol{H}(L)$ with $\hat{\boldsymbol{\varphi}}_k(t-i)$, and the estimates $\hat{v}_k(t-i)$ of $v(t-i)$ at iteration k can be computed by (26).

Replacing $\boldsymbol{H}(L)$ in (27) with $\hat{\boldsymbol{H}}_k(L)$, we can obtain the least squares based iterative identification algorithm for estimating $\boldsymbol{\Theta}$ of the CARMA systems as follows [21]:

$$\hat{\boldsymbol{\Theta}}_k = [\hat{\boldsymbol{H}}_k^{\mathrm{T}}(L)\hat{\boldsymbol{H}}_k(L)]^{-1}\hat{\boldsymbol{H}}_k^{\mathrm{T}}(L)\boldsymbol{Y}(L), \tag{28}$$

$$\boldsymbol{Y}(L) = [y(1), y(2), \ldots, y(L)]^{\mathrm{T}}, \tag{29}$$

$$\hat{\boldsymbol{H}}_k(L) = [\hat{\boldsymbol{\varphi}}_k(1), \hat{\boldsymbol{\varphi}}_k(2), \ldots, \hat{\boldsymbol{\varphi}}_k(L)]^{\mathrm{T}}, \tag{30}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_k(t) = [&-y(t-1), -y(t-2), \ldots, -y(t-n_a), u(t-1), u(t-2), \ldots, u(t-n_b),\\ &\hat{v}_{k-1}(t-1), \hat{v}_{k-1}(t-2), \ldots, \hat{v}_{k-1}(t-n_d)]^{\mathrm{T}},\end{aligned} \tag{31}$$

$$\hat{v}_k(t) = y(t) - \hat{\boldsymbol{\varphi}}_k^{\mathrm{T}}(t)\hat{\boldsymbol{\Theta}}_k. \tag{32}$$

Compared the LSI algorithm and the two-stage LSI algorithm, we can draw the following conclusions.

Table 1
The parameter estimates and their errors with $\sigma^2 = 0.10^2$.

Algorithms	$t = L$	a_1	a_2	b_1	b_2	d_1	δ (%)
LSI	1000	−1.59746	0.79877	0.39895	0.30981	−0.59785	2.20821
	2000	−1.60034	0.80072	0.39985	0.30210	−0.61364	1.34686
	3000	−1.60041	0.80051	0.40074	0.29990	−0.62064	0.98659
Two-stage LSI	1000	−1.60215	0.80342	0.39885	0.30721	−0.59985	2.08739
	2000	−1.60263	0.80298	0.39983	0.30109	−0.61427	1.32626
	3000	−1.60170	0.80203	0.40072	0.29983	−0.62067	0.99376
True values		−1.6000	0.8000	0.4000	0.3000	−0.6400	

Table 2
The parameter estimates and their errors with $\sigma^2 = 0.50^2$.

Algorithms	$t = L$	a_1	a_2	b_1	b_2	d_1	δ (%)
LSI	1000	−1.58859	0.79212	0.39474	0.34862	−0.58797	3.70276
	2000	−1.59951	0.79752	0.39930	0.31149	−0.61263	1.51670
	3000	−1.60004	0.79577	0.40384	0.30051	−0.62025	1.04672
Two-stage LSI	1000	−1.60651	0.80940	0.39447	0.33795	−0.60288	2.77809
	2000	−1.60739	0.80748	0.39914	0.30773	−0.61841	1.28484
	3000	−1.60382	0.80427	0.40353	0.30097	−0.62248	0.95681
True values		−1.6000	0.8000	0.4000	0.3000	−0.6400	

Table 3
The parameter estimates and their errors with $\sigma^2 = 1.00^2$.

Algorithms	$t = L$	a_1	a_2	b_1	b_2	d_1	δ (%)
LSI	1000	−1.58136	0.78563	0.38961	0.39558	−0.57972	5.89917
	2000	−1.59672	0.79103	0.39868	0.32381	−0.60977	2.01938
	3000	−1.59830	0.78690	0.40799	0.30200	−0.61845	1.35311
Two-stage LSI	1000	−1.60536	0.80536	0.38953	0.37929	−0.60125	4.54018
	2000	−1.60376	0.80163	0.39839	0.31993	−0.61456	1.66025
	3000	−1.59906	0.79814	0.40711	0.30542	−0.61766	1.22932
True values		−1.6000	0.8000	0.4000	0.3000	−0.6400	

In the LSI algorithm, the dimension of the covariance matrix $S_3 := [\hat{\boldsymbol{H}}_k^{\mathrm{T}}(L)\hat{\boldsymbol{H}}_k(L)]^{-1}$ is $(n_a + n_b + n_d) \times (n_a + n_b + n_d)$ which is larger than those of S_1 and S_2 and thus the two-stage LSI algorithm has a high computational efficiency.

4. Example

Consider the following simulation system:

$$
\begin{aligned}
&A(z)y(t) = B(z)u(t) + D(z)v(t),\\
&A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} = 1 - 1.60z^{-1} + 0.80z^{-2},\\
&B(z) = b_1 z^{-1} + b_2 z^{-2} = 0.40z^{-1} + 0.30z^{-2},\\
&D(z) = 1 + d_1 z^{-1} = 1 - 0.64z^{-1},\\
&\boldsymbol{\theta} = [a_1, a_2, b_1, b_2]^{\mathrm{T}} = [-0.60, 0.80, 0.40, 0.30]^{\mathrm{T}},\\
&\boldsymbol{\vartheta} = d_1 = -0.64,\\
&\boldsymbol{\Theta} = [\boldsymbol{\theta}^{\mathrm{T}}, \boldsymbol{\vartheta}^{\mathrm{T}}]^{\mathrm{T}}.
\end{aligned}
$$

In simulation, $\{u(t)\}$ is taken as an uncorrelated persistent excitation signal sequence with zero mean and unit variance, $\{v(t)\}$ as a white noise sequence with zero mean and variance σ^2. Applying the LSI and the two-stage LSI algorithm to estimate the parameters of this system, the parameter estimates with different data length and noise variances are shown in Tables 1–3. The estimation errors for the two-stage LSI algorithm versus k are shown in Tables 4–6 and Figs. 3–5 where $\delta := \|\hat{\boldsymbol{\Theta}}_k - \boldsymbol{\Theta}\|/\|\boldsymbol{\Theta}\|$. When $\sigma^2 = 0.10^2$, $\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$, the corresponding noise-to-signal ratios are $\delta_{\mathrm{ns}} = 7.66\%$, $\delta_{\mathrm{ns}} = 38.2\%$, and $\delta_{\mathrm{ns}} = 76.5\%$, respectively.

From the simulation results in Tables 1–6 and Figs. 3–5, we can draw the following conclusions.

- The estimation errors δ given by the LSI algorithm and the two-stage LSI algorithm become smaller as the data length L increases.
- Under the same data length and noise variance, the estimation accuracy of the two algorithms is close.
- As the noise-to-signal ratio becomes small, the parameter estimation errors given by the two-stage LSI algorithm become small and the parameter estimates converge fast to their true values for the same data length $t = L$.
- The two-stage LSI algorithm has fast convergence rates and needs only a few iterations to converge to their true values.

Table 4
The two-stage LSI estimates and errors versus iteration k ($L = 1000$).

σ^2	k	a_1	a_2	b_1	b_2	d_1	δ (%)
0.10^2	1	−1.59185	0.79393	0.39630	0.31191	−0.57976	3.17396
	2	−1.60180	0.80310	0.39877	0.30737	−0.59918	2.12038
	5	−1.60215	0.80342	0.39885	0.30721	−0.59985	2.08740
	10	−1.60215	0.80342	0.39885	0.30721	−0.59985	2.08739
0.50^2	1	−1.49095	0.70808	0.38465	0.38643	−0.46948	12.16564
	2	−1.58094	0.78698	0.39230	0.34868	−0.57337	4.37908
	5	−1.60623	0.80916	0.39445	0.33806	−0.60257	2.78967
	10	−1.60651	0.80940	0.39447	0.33795	−0.60288	2.77809
1.00^2	1	−1.36445	0.60223	0.37486	0.47944	−0.34398	23.60467
	2	−1.50232	0.71847	0.38326	0.42213	−0.49121	11.77593
	5	−1.59739	0.79863	0.38905	0.38261	−0.59274	4.87865
	10	−1.60536	0.80536	0.38953	0.37929	−0.60125	4.54018
True values		−1.6000	0.8000	0.4000	0.3000	−0.6400	

Table 5
The two-stage LSI estimates and errors versus iteration k ($L = 2000$).

σ^2	k	a_1	a_2	b_1	b_2	d_1	δ (%)
0.10^2	1	−1.59450	0.79572	0.39902	0.30434	−0.60283	1.93844
	2	−1.60248	0.80284	0.39981	0.30115	−0.61406	1.33561
	5	−1.60263	0.80298	0.39983	0.30109	−0.61427	1.32626
	10	−1.60263	0.80298	0.39983	0.30109	−0.61427	1.32626
0.50^2	1	−1.49354	0.70788	0.39515	0.35248	−0.49772	10.53550
	2	−1.58517	0.78804	0.39836	0.31646	−0.59486	2.63224
	5	−1.60723	0.80734	0.39914	0.30780	−0.61824	1.28892
	10	−1.60739	0.80748	0.39914	0.30773	−0.61841	1.28484
1.00^2	1	−1.35502	0.59085	0.39167	0.41784	−0.36110	22.50624
	2	−1.50120	0.71473	0.39562	0.36030	−0.51005	9.86683
	5	−1.59664	0.79560	0.39820	0.32274	−0.60731	2.04856
	10	−1.60376	0.80163	0.39839	0.31993	−0.61456	1.66025
True values		−1.6000	0.8000	0.4000	0.3000	−0.6400	

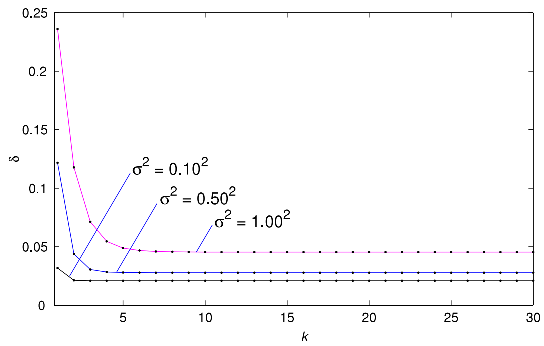

Fig. 3. The two-stage LSI estimation errors δ versus k with different σ^2($L = 1000$).

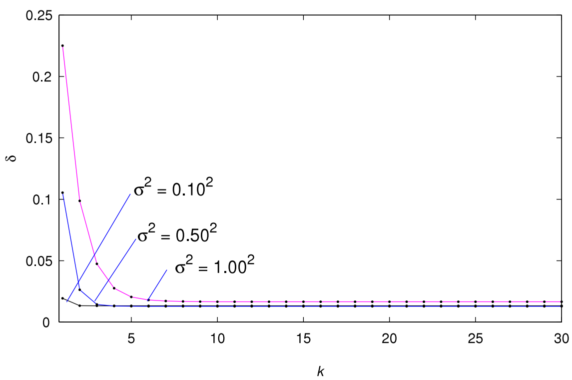

Fig. 4. The two-stage LSI estimation errors δ versus k with different σ^2 ($L = 2000$).

Table 6
The two-stage LSI estimation and errors versus iteration k ($L = 3000$).

σ^2	k	a_1	a_2	b_1	b_2	d_1	δ (%)
0.10^2	1	−1.59450	0.79548	0.40011	0.30229	−0.61184	1.48298
	2	−1.60159	0.80194	0.40071	0.29987	−0.62055	0.99915
	5	−1.60170	0.80203	0.40072	0.29983	−0.62067	0.99376
	10	−1.60170	0.80203	0.40072	0.29983	−0.62067	0.99376
0.50^2	1	−1.49087	0.70455	0.40120	0.34475	−0.50657	10.28509
	2	−1.58279	0.78570	0.40309	0.30912	−0.60089	2.34527
	5	−1.60369	0.80415	0.40353	0.30103	−0.62234	0.96080
	10	−1.60382	0.80427	0.40353	0.30097	−0.62248	0.95681
1.00^2	1	−1.34827	0.58382	0.40465	0.40543	−0.36604	22.55197
	2	−1.49692	0.71085	0.40611	0.34616	−0.51518	9.70069
	5	−1.59223	0.79230	0.40704	0.30815	−0.61080	1.67920
	10	−1.59906	0.79814	0.40711	0.30542	−0.61766	1.22932
True values		−1.6000	0.8000	0.4000	0.3000	−0.6400	

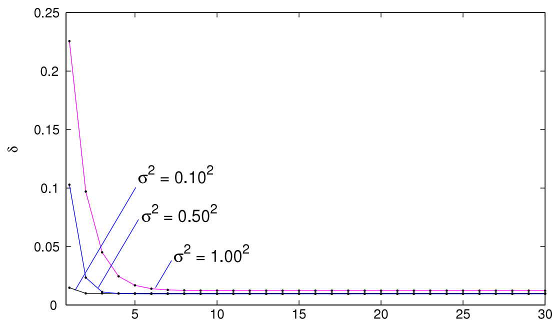

Fig. 5. The two-stage LSI estimation errors δ versus k with different σ^2($L = 3000$).

5. Conclusions

This paper presents a two-stage LSI algorithm for CARMA models. The proposed algorithm requires less computational load than the least squares based iterative algorithm. The simulation results show that the proposed algorithm has fast convergence rates and can generate highly accurate parameter estimates after only several iterations.

References

[1] M. Dehghan, M. Hajarian, Matrix equations over (R,S)-symmetric and (R,S)-skew symmetric matrices, Computers & Mathematics with Applications 59 (11) (2010) 3583–3594.
[2] M. Dehghan, M. Hajarian, An iterative algorithm for solving a pair of matrix equations AYB=E, CYD=F over generalized centro-symmetric matrices, Computers & Mathematics with Applications 56 (12) (2008) 3246–3260.
[3] M. Dehghan, M. Hajarian, An iterative method for solving the generalized coupled Sylvester matrix equations over generalized bisymmetric matrices, Applied Mathematical Modelling 34 (3) (2010) 639–654.
[4] M. Dehghan, M. Hajarian, Finite iterative algorithms for the reflexive and anti-reflexive solutions of the matrix equation A1X1B1 + A2X2B2 = C, Mathematical and Computer Modelling 49 (9–10) (2009) 1937–1959.
[5] L. Xie, J. Ding, F. Ding, Gradient based iterative solutions for general linear matrix equations, Computers & Mathematics with Applications 58 (7) (2009) 1441–1448.
[6] J. Ding, Y.J. Liu, F. Ding, Iterative solutions to matrix equations of form AiXBi = Fi, Computers & Mathematics with Applications 59 (11) (2010) 3500–3507.
[7] F. Ding, T. Chen, Gradient based iterative algorithms for solving a class of matrix equations, IEEE Transactions on Automatic Control 50 (8) (2005) 1216–1221.
[8] L. Xie, J. Ding, F. Ding, Gradient based iterative solutions for general linear matrix equations, Computers and Mathematics with Applications 58 (7) (2009) 1441–1448.
[9] A.G. Wu, B. Li, Y. Yang, G.R. Duan, Finite iterative solutions to coupled Sylvester-conjugate matrix equations, Applied Mathematical Modelling 35 (3) (2011) 1065–1080.
[10] Y. Shi, H. Fang, Kalman filter based identification for systems with randomly missing measurements in a network environment, International Journal of Control 83 (3) (2010) 538–551.
[11] Y. Shi, H. Fang, M. Yan, Kalman filter based adaptive control for networked systems with unknown parameters and randomly missing outputs, International Journal of Robust and Nonlinear Control 19 (18) (2009) 1976–1992. (special issue on control with limited information (part II)).
[12] B. Yu, Y. Shi, H. Huang, l–2 and *l*-infnity filtering for multirate systems using lifted models, Circuits, Systems, and Signal Processing 27 (5) (2008) 699–711.
[13] H. Fang, J. Wu, Y. Shi, Genetic adaptive state estimation with missing input/output data, Proceedings of the Institution of Mechanical Engineers-Part I: Journal of Systems and Control Engineering 224 (5) (2010) 611–617.
[14] Y. Shi, F. Ding, T. Chen, 2-norm based recursive design of transmultiplexers with designable filter length, Circuits, Systems and Signal Processing 25 (4) (2006) 447–462.
[15] Y.S. Xiao, Y. Zhang, J. Ding, J.Y. Dai, The residual based interactive least squares algorithms and simulation studies, Computers & Mathematics with Applications 58 (6) (2009) 1190–1197.

[16] L.Y. Wang, L. Xie, X.F. Wang, The residual based interactive stochastic gradient algorithms for controlled moving average models, Applied Mathematics and Computation 211 (2) (2009) 442–449.
[17] F. Ding, T. Chen, L. Qiu, Bias compensation based recursive least squares identification algorithm for MISO systems, IEEE Transactions on Circuits and Systems - II: Express Briefs 53 (5) (2006) 349–353.
[18] F. Ding, Y. Shi, T. Chen, Performance analysis of estimation algorithms of non-stationary ARMA processes, IEEE Transactions on Signal Processing 54 (3) (2006) 1041–1053.
[19] F. Ding, P.X. Liu, G.L. Liu, Gradient based and least-squares based iterative identification methods for OE and OEMA systems, Digital Signal Processing 20 (3) (2010) 664–677.
[20] D.Q. Wang, G.W. Yang, R.F. Ding, Gradient-based iterative parameter estimation for Box–Jenkins systems, Computers & Mathematics with Applications 60 (5) (2010) 1200–1208.
[21] Y.J. Liu, D.Q. Wang, F. Ding, Least squares based iterative identification for identifying Box–Jenkins models with finite measurement date, Digital Signal Processing 20 (5) (2010) 1458–1467.
[22] F. Ding, T. Chen, Identification of Hammerstein nonlinear ARMAX systems, Automatica 41 (9) (2005) 1479–1489.
[23] H.Q. Han, L. Xie, F. Ding, X.G. Liu, Hierarchical least-squares based iterative identification for multivariable systems with moving average noises, Automatica 51 (9–10) (2010) 1213–1220.
[24] Z.N. Zhang, F. Ding, X.G. Liu, Hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems, Computers & Mathematics with Applications 61 (3) (2011) 672–682.
[25] B. Bao, Y.Q. Xu, J. Sheng, R.F. Ding, Least squares based iterative parameter estimation algorithm for multivariable controlled ARMA system modelling with finite measurement data, Mathematical and Computer Modelling 53 (9–10) (2011) 1664–1669.
[26] F. Ding, Y.J. Liu, B. Bao, Gradient based and least squares based iterative estimation algorithms for multi-input multi-output systems, Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering (2011), in press (doi:10.1177/0959651811409491).
[27] E.W. Bai, An optimal two-stage identification algorithm for Hammerstein–Wiener nonlinear systems, Automatica 34 (3) (1998) 333–338.
[28] K. Li, J.X. Peng, E. Bai, A two-stage algorithm for identification of nonlinear dynamic systems, Automatica 42 (7) (2006) 1189–1197.
[29] S.H. Hwang, S.T. Lai, Use of two-stage least-squares algorithms for identification of continuous systems with time delay based on pulse responses, Automatica 40 (9) (2004) 1561–1568.
[30] Z.H. Cao, Rounding error analysis of two-stage iterative methods for large linear systems, Applied Mathematics and Computation 139 (2–3) (2003) 371–381.
[31] H.H. Duan, J. Jia, R.F. Ding, Two-stage recursive least squares parameter estimation algorithm for output error models, Mathematical and Computer Modelling 55 (x) (2012) doi:10.1016/j.mcm.2011.09.039.
[32] D.Q. Wang, F. Ding, Extended stochastic gradient identification algorithms for Hammerstein–Wiener ARMAX systems, Computers & Mathematics with Applications 56 (12) (2008) 3157–3164.
[33] J. Chen, Y. Zhang, R.F. Ding, Auxiliary model based multi-innovation algorithms for multivariable nonlinear systems, Mathematical and Computer Modelling 52 (9–10) (2010) 1428–1434.
[34] D.Q. Wang, Least squares-based recursive and iterative estimation for output error moving average (OEMA) systems using data filtering, IET Control Theory and Applications 5 (14) (2011) 1648–1657.
[35] D.Q. Wang, F. Ding, Input–output data filtering based recursive least squares parameter estimation for CARARMA systems, Digital Signal Processing 20 (4) (2010) 991–999.
[36] L.L. Han, F. Ding, Parameter estimation for multirate multi-input systems using auxiliary model and multi-innovation, Journal of Systems Engineering and Electronics 21 (6) (2010) 1079–1083.
[37] L.L. Han, F. Ding, Identification for multirate multi-input systems using the multi-innovation identification theory, Computers & Mathematics with Applications 57 (9) (2009) 1438–1449.
[38] F. Ding, T. Chen, Combined parameter and output estimation of dual-rate systems using an auxiliary model, Automatica 40 (10) (2004) 1739–1748.
[39] F. Ding, T. Chen, Identification of dual-rate systems based on finite impulse response models, International Journal of Adaptive Control and Signal Processing 18 (7) (2004) 589–598.
[40] F. Ding, T. Chen, Least squares based self-tuning control of dual-rate systems, International Journal of Adaptive Control and Signal Processing 18 (8) (2004) 697–714.
[41] F. Ding, T. Chen, Parameter estimation of dual-rate stochastic systems by using an output error method, IEEE Transactions on Automatic Control 50 (9) (2005) 1436–1441.
[42] Y.J. Liu, J. Sheng, R.F. Ding, Convergence of stochastic gradient estimation algorithm for multivariable ARX-like systems, Computers & Mathematics with Applications 59 (8) (2010) 2615–2627.
[43] F. Ding, T. Chen, Performance analysis of multi-innovation gradient type identification methods, Automatica 43 (1) (2007) 1–14.
[44] Y.J. Liu, Y.S. Xiao, X.L. Zhao, Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model, Applied Mathematics and Computation 215 (4) (2009) 1477–1483.
[45] F. Ding, P.X. Liu, G. Liu, Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises, Signal Processing 89 (10) (2009) 1883–1890.
[46] J.B. Zhang, F. Ding, Y. Shi, Self-tuning control based on multi-innovation stochastic gradient parameter estimation, Systems & Control Letters 58 (1) (2009) 69–75.
[47] F. Ding, P.X. Liu, G. Liu, Multi-innovation least squares identification for system modeling, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 40 (3) (2010) 767–778.
[48] F. Ding, Several multi-innovation identification methods, Digital Signal Processing 20 (4) (2010) 1027–1039.
[49] D.Q. Wang, F. Ding, Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems, Digital Signal Processing 20 (3) (2010) 750–762.
[50] L. Xie, H.Z. Yang, F. Ding, Modeling and identification for non-uniformly periodically sampled-data systems, IET Control Theory & Applications 4 (5) (2010) 784–794.
[51] L.L. Han, F. Ding, Multi-innovation stochastic gradient algorithms for multi-input multi-output systems, Digital Signal Processing 19 (4) (2009) 545–554.
[52] Y.J. Liu, L. Yu, F. Ding, Multi-innovation extended stochastic gradient algorithm and its performance analysis, Circuits, Systems and Signal Processing 29 (4) (2010) 649–667.
[53] F. Ding, T. Chen, Hierarchical gradient-based identification of multivariable discrete-time systems, Automatica 41 (2) (2005) 315–325.
[54] F. Ding, T. Chen, Hierarchical least squares identification methods for multivariable systems, IEEE Transactions on Automatic Control 50 (3) (2005) 397–402.
[55] F. Ding, T. Chen, Hierarchical identification of lifted state–space models for general dual-rate systems, IEEE Transactions on Circuits and Systems—I: Regular Papers 52 (6) (2005) 1179–1187.
[56] F. Ding, T. Chen, A gradient based adaptive control algorithm for dual-rate systems, Asian Journal of Control 8 (4) (2006) 314–323.
[57] F. Ding, T. Chen, Z. Iwai, Adaptive digital control of Hammerstein nonlinear systems with limited output sampling, SIAM Journal on Control and Optimization 45 (6) (2007) 2257–2276.
[58] F. Ding, L. Qiu, T. Chen, Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems, Automatica 45 (2) (2009) 324–332.
[59] Y.J. Liu, L. Xie, F. Ding, An auxiliary model based recursive least squares parameter estimation algorithm for non-uniformly sampled multirate systems, Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering 223 (4) (2009) 445–454.

[60] F. Ding, G. Liu, X.P. Liu, Partially coupled stochastic gradient identification methods for non-uniformly sampled systems, IEEE Transactions on Automatic Control 55 (8) (2010) 1976–1981.
[61] F. Ding, J. Ding, Least squares parameter estimation with irregularly missing data, International Journal of Adaptive Control and Signal Processing 24 (7) (2010) 540–553.
[62] F. Ding, G. Liu, X.P. Liu, Parameter estimation with scarce measurements, Automatica 47 (8) (2011) 1646–1655.
[63] L. Xie, H.Z. Yang, F. Ding, Recursive least squares parameter estimation for non-uniformly sampled systems based on the data filtering, Mathematical and Computer Modelling 54 (1–2) (2011) 315–324.
[64] J. Chen, F. Ding, Modified stochastic gradient algorithms with fast convergence rates, Journal of Vibration and Control 17 (9) (2011) 1281–1286.
[65] J. Ding, Y. Shi, H.G. Wang, F. Ding, A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems, Digital Signal Processing 20 (4) (2010) 1238–1249.
[66] F. Ding, P.X. Liu, H.Z. Yang, Parameter identification and intersample output estimation for dual-rate systems, IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans 38 (4) (2008) 966–975.
[67] F. Ding, H.Z. Yang, F. Liu, Performance analysis of stochastic gradient algorithms under weak conditions, Science in China Series F: Information Sciences 51 (9) (2008) 1269–1280.
[68] F. Ding, T. Chen, Performance bounds of the forgetting factor least squares algorithm for time-varying systems with finite measurement data, IEEE Transactions on Circuits and Systems—I: Regular Papers 52 (3) (2005) 555–566.
[69] W. Wang, F. Ding, J.Y. Dai, Maximum likelihood least squares identification for systems with autoregressive moving average noise, Applied Mathematical Modelling 36 (x) (2012) doi:10.1016/j.apm.2011.07.083.
[70] J. Ding, F. Ding, X.P. Liu, G. Liu, Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data, IEEE Transactions on Automatic Control 56 (11) (2011) 2677–2683.
[71] J. Ding, F. Ding, Bias compensation based parameter estimation for output error moving average systems, International Journal of Adaptive Control and Signal Processing 25 (12) (2011) 1100–1111.
[72] L.F. Zhuang, F. Pan, F. Ding, Parameter and state estimation algorithm for single-input single-output linear systems using the canonical state space models, Applied Mathematical Modelling 26 (x) (2012) doi:10.1016/j.apm.2011.10.028.
[73] J.H. Li, F. Ding, Maximum likelihood stochastic gradient estimation for Hammerstein systems with colored noise based on the key term separation technique, Computers & Mathematics with Applications 62 (11) (2011) 4170–4177.
[74] F. Ding, P.X. Liu, G. Liu, Identification methods for Hammerstein nonlinear systems, Digital Signal Processing 21 (2) (2011) 215–238.
[75] D.Q. Wang, Y.Y. Chu, F. Ding, Auxiliary model-based RELS and MI–ELS algorithms for Hammerstein OEMA systems, Computers & Mathematics with Applications 59 (9) (2010) 3092–3098.
[76] D.Q. Wang, Y.Y. Chu, G.W. Yang, F. Ding, Auxiliary model-based recursive generalized least squares parameter estimation for Hammerstein OEAR systems, Mathematical and Computer Modelling 52 (1–2) (2010) 309–317.
[77] D.Q. Wang, F. Ding, Least squares based and gradient based iterative identification for Wiener nonlinear systems, Signal Processing 91 (5) (2011) 1182–1189.
[78] F. Ding, Y. Shi, T. Chen, Auxiliary model based least-squares identification methods for Hammerstein output-error systems, Systems & Control Letters 56 (5) (2007) 373–380.
[79] F. Ding, Y. Shi, T. Chen, Gradient-based identification methods for Hammerstein nonlinear ARMAX models, Nonlinear Dynamics 45 (1–2) (2006) 31–43.
[80] D.Q. Wang, F. Ding, Extended stochastic gradient identification algorithms for Hammerstein–Wiener ARMAX systems, Computers & Mathematics with Applications 56 (12) (2008) 3157–3164.